

**Initial Conditions and Heterogeneity in Cross-Country Growth:
An Iterative Bayesian Model Averaging (IBMA) Analysis**

Jeffrey Begun
Department of Economics
University of Washington

April, 2009

Abstract: This paper tests for parameter heterogeneity in cross-country growth determinants due to threshold effects in initial income and human capital, and allows for model uncertainty in growth determinants through use of Iterative Bayesian Model Averaging (IBMA, a procedure which runs BMA recursively on subsets of data and thus allows for a large number of candidate regressors to be processed efficiently). Allowing for model uncertainty in cross-country growth studies is crucial since while scores of theory-supported growth determinants have been proposed, the “true” underlying growth regression is not known. Further, while several authors have found evidence of multiple regimes due to initial conditions, and some have allowed for model uncertainty, they have only tested for parameter heterogeneity in the original Solow variables or in a limited set of explanatory variables (thus not allowing for heterogeneous growth in many potential covariates, and not making the results fully comparable to any variable-rich cross-country growth dataset).

This study builds on past work by allowing for model uncertainty and parameter heterogeneity in a full suite of candidate regressors from a well-known cross-country growth dataset. Initial income and human capital (public education share) are found to be important threshold variables for growth heterogeneity, which lends supports to Durlauf and Johnson’s (1995) and Masanjala and Papageorgiou’s (2004) findings but contradicts Crespo Cuaresma and Doppelhofer (2007) who also employ Bayesian techniques.

I. Introduction

Although a number of cross-country growth studies such as Barro (1991) and Sala-i-Martin, Doppelhofer, and Miller (2004) assume a homogeneous growth process, a large body of theoretical (e.g., Azariadis and Drazen, 1990; Durlauf, 1993; Galor and Zeria, 1993) and empirical work (e.g., Durlauf and Johnson, 1995; Liu and Stengos, 1999; Hansen, 2000) supports the possibility that countries do not all follow the same growth path. There are several theoretical reasons why countries may grow in a heterogeneous way: individual growth determinants such as physical and human capital may have disparate effects on countries' economic growth rates due to threshold effects, non-standard production functions, or nonlinearities such as increasing returns. Since countries often possess vastly different economic, institutional, and social characteristics, these theoretical possibilities are likely to hold in practice and imposing homogeneity restrictions on all potential growth determinants may be unwise.¹

While the problem of cross-country growth heterogeneity has been approached from a variety of angles, one major branch of the literature has tried to identify subsets of countries that follow similar laws of motion for growth. This paper follows studies such as Durlauf and Johnson (1995), Hansen (2000), and Masanjala and Papageorgiou (2004) that employ initial income and human capital as thresholds in order to test for multiple regimes. As stressed in Brock and Durlauf (2001), however, there are two types of model uncertainty that should be addressed in cross-country growth studies: heterogeneity uncertainty, and theory uncertainty (i.e. which determinants really matter for growth). The current study takes the approach of Brock and Durlauf (2001) and Crespo Cuaresma and Doppelhofer (2007) in that it tests for both.

¹ See Durlauf, Johnson, and Temple (2005) for a more in-depth discussion.

Specifically, it uses Iterative Bayesian Model Averaging (IBMA), which applies Bayesian Model Averaging (BMA) recursively, to allow for model uncertainty.²

The current paper relates to Crespo Cuaresma and Doppelhofer's 2007 work on Bayesian Averaging of Thresholds (BAT) in that both studies focus on initial conditions and use Bayesian techniques to allow for model uncertainty, but there are several key differences. First, this study uses initial income and human capital as thresholds, both individually and as a combined threshold, unlike Crespo Cuaresma and Doppelhofer who test initial income and openness on a separate basis only. Thus, the results from this paper are more comparable to the benchmark Durlauf and Johnson (1995) study. In addition, Crespo Cuaresma and Doppelhofer only test the 21 (out of 67) variables in their dataset which were found to be robust in Sala-i-Martin *et al.* (2004), which assumed homogenous growth. This paper, however, tests for parameter heterogeneity in all 41 candidate regressors from the benchmark Fernández, Ley, and Steel (2001) dataset.³ This is important because there may be variables which are deemed to be non-robustly related to growth when the growth process is assumed to be homogenous, but which may contain explanatory power for growth in a subsample of countries. Testing the full suite of variables for heterogeneity, which is possible with IBMA, also allows for a more straightforward comparison with the original model averaging results from Fernández *et al.*

This study finds that initial income and public education as a percent of GDP are both important threshold variables for identifying multiple regimes. IBMA exposes a great deal of heterogeneity in the growth process as only around half of the variables found to be important for growth in the global sample are found to matter in the initial GDP and public education share

² See Sections III-V for a detailed discussion of BMA and IBMA.

³ Between three and five variables in each specification were not tested for heterogeneity since their interaction terms for high initial income/human capital were either perfectly (or extremely highly) collinear or had less than two non-zero observations. See the notes to Tables 1-5 for a list of which interaction terms were dropped in each case.

subsamples. Furthermore, there are considerable differences as to which variables matter for growth in the low/high income subsamples and the low/high public education share subsamples. Specifically, 9 (7) variables which are found to be effective for growth in one of the initial GDP (public education share) subsamples are not effective in the other subsample. Several variables also change signs between the global specification and the subsamples, and within the subsamples, which suggests that even covariates which are found to be robustly related to growth in different groups of countries have disparate and perhaps surprising effects due to varying initial conditions. Finally, primary schooling on its own does not appear to be a key threshold variable, though when combined with initial income as a group threshold it does uncover evidence of parameter heterogeneity.

In sum, the present study, which analyses the largest pool of potential growth determinants to date when testing for multiple regimes due to initial conditions, uncovers a vast amount of heterogeneity in the growth process. The results lend support to studies such as Durlauf and Johnson (1995), Hansen (2000), and Crespo Cuaresma (2002) which also find that initial income and human capital are important threshold variables. On the other hand, the results run counter to Crespo Cuaresma and Doppelhofer (2007) who find that initial income is not an important threshold variable that leads to heterogeneity in cross-country growth determinants.

II. Cross-Country Heterogeneity

Since cross-country growth studies often include more than 80 countries with vastly different economic, political, institutional, religious, geographic, and demographic properties, imposing homogeneity restrictions on all potential growth determinants may not be wise. Brock and Durlauf (2001) note that although all socioeconomic datasets are subject to similar homogeneity

concerns, these concerns are particularly salient when comparing vastly different entities such as countries.

While the basic Solow model does not allow for heterogeneity, a number of theoretical models such as those in Azariadis and Drazen (1990) and Durlauf (1993) deliver multiple steady states. These papers stress the importance of threshold effects and nonlinearities in the growth process due to human capital (Azariadis and Drazen) and technological externalities (Durlauf; Azariadis and Drazen) which result from social returns to scale as individual economic agents amass human capital and technological know-how. In addition, heterogeneity may result from non-standard production functions. Identical Cobb-Douglas technology, which is required for the linear Solow growth model which lies at the heart of many cross-country growth studies, may be a dubious assumption. For example, Mansala and Papageorgiou (2004) find that the more general Constant-Elasticity-of-Substitution specification is more appropriate in a cross-country growth context than the standard Cobb-Douglas model.

In light of these theoretical concerns, Durlauf and Johnson (1995) test for and find evidence of multiple regimes by dividing the countries in their sample into clusters based on differences in initial income and literacy. Since Durlauf and Johnson a large number of authors such as Liu and Stengos (1999), Durlauf, Kourtellos, and Minkin (2001), Masanjala and Papageorgiou (2004), and Ardiç (2006) have allowed for heterogeneity in the basic Solow growth determinants and found evidence that countries grow differently.⁴

III. Allowing for Model Uncertainty

⁴ Papageorgiou (2002) also finds evidence of heterogeneity in the basic Solow variables when splitting the main sample into subsamples based on initial income, initial literacy, and trade share.

Although Durlauf and Johnson (1995) and subsequent work shed much light on heterogeneity in the cross-country growth context, these studies typically focus on the Solow growth model or on the basic Solow specification with a few additional covariates. Starting with early cross-country growth regressions such as Kormendi and Meguire (1985) and Barro (1991), however, researchers have found evidence that a wide variety of variables have explanatory power. Durlauf *et al.* (2005) note that hundreds of growth determinants have been identified by researchers, more determinants than there are countries in the world.

Given the large number of potential regressors and the relatively small set of countries in typical cross-country datasets, the researcher who wishes to employ OLS is left with the unenviable task of deciding which subset of variables to include in his or her specification, and which to discard. Given the fact that there is often little theoretical justification for including certain variables while leaving others out⁵, different authors can come to reasonable yet contradictory conclusions about the importance of various growth determinants. In order to try to overcome this type of difficulty, Leamer (1978) stressed that empirical work should allow for model uncertainty. His 1983 paper on Extreme Bounds Analysis (EBA) provides a mechanism for doing this by testing all possible models and then deeming regressors which switched signs to be non-robust. Critics pointed out that this technique was biased towards returning “too few” robust regressors and that it was not grounded in well-developed statistical theory (Sala-i-Martin, 1997; Sala-i-Martin *et al.*, 2004).

Due to the shortcomings of EBA and related methods such as the one employed in Sala-i-Martin (1997), researchers have increasingly turned to Bayesian Model Averaging (BMA) when faced with an unwieldy mass of potential regressors. BMA was first brought to economics by

⁵ Brock and Durlauf (2001) describe this as the “open-endedness” problem since the validity of one causal theory of growth does not invalidate other growth theories.

Multon (1991) and Palm and Zeller (1992), is featured in the oft-cited Fernández, Ley, and Steel (2001) study, and has gained prominence in empirical economics as of late as computing speed has rapidly increased. BMA has several advantages over EBA since it is firmly grounded in statistical foundations and it can find the “best” model by testing any subset of variables. BMA has also been shown to deliver better out-of-sample predictions than predictions based on a single model (Raftery, 1995; Raftery *et al.*, 1997).⁶

Several studies of growth heterogeneity such as Brock and Durlauf (2001), Crespo Cuaresma and Doppelhofer (2007), Eicher, Papageorgiou, and Roehn (2007), and Begun (2008) use Bayesian methods to test for cross-country growth heterogeneity. Like Durlauf and Johnson (1995) and Masanjala and Papageorgiou (2004), Crespo Cuaresma and Doppelhofer (2007) also focus on parameter heterogeneity due to initial conditions in their Bayesian Averaging of Thresholds (BAT) paper. Their approach has an advantage, however, in that it uses model averaging to allow for heterogeneity in a larger suite of 21 regressors. Unfortunately, their analysis is not fully comparable since Durlauf and Johnson (1995) and Masanjala and Papageorgiou (2004) split their sample based on initial income and human capital, while Crespo Cuaresma and Doppelhofer test for subsample heterogeneity due to differences in initial income and openness. Also, unlike Durlauf and Johnson (1995) they do not test for threshold interactions.

The current study builds off of Durlauf and Johnson (1995) and Crespo Cuaresma and Doppelhofer (2007). Like Durlauf and Johnson it tests for threshold effects in initial income and human capital, and also tests for parameter heterogeneity with a combined high-income/high-human-capital threshold. Like Crespo Cuaresma and Doppelhofer it employs Bayesian

⁶ See Raftery (1995), Draper (1995), and Hoeting, Madigan, Raftery, and Volinsky (1999) for a more in-depth discussion of BMA.

techniques and model averaging in the analysis. However, unlike Crespo Cuaresma and Doppelhofer, who only test the 21 variables that Sala-i-Martin *et al.* (2004) find to be robustly related to growth (out of 67), this study uses IBMA to test all 41 candidate regressors in the Fernández, Ley, and Steel (2001) dataset for parameter heterogeneity.⁷ This is important because there may be potential explanatory variables which are found to be non-robust under an assumption of homogeneity, but which may be robustly related to growth in a given subsample. Testing the full suite of variables, which is possible with IBMA, also allows for a more straightforward comparison with the original Fernández *et al.* model averaging results under parameter homogeneity.

IV. Bayesian Model Averaging

This study uses Iterative Bayesian Model Averaging, an offshoot of BMA that performs BMA multiple times on subsets of the data, in order to gain inference.⁸ Unlike some model averaging techniques, BMA is able to average over all possible models, where the posterior model probability is used to weigh each model. The posterior model probability for any given model equals the conditional probability that the model is the true model after all relevant data has been accounted for. These posterior probabilities are calculated by incorporating the likelihood function and prior probabilities selected by the researcher into Bayes' theorem. The rest of this section presents a more formal model of BMA as used in this paper.

⁷ The current study also tested the Sala-i-Martin *et al.* (2004) dataset for growth heterogeneity using IBMA. The results are qualitatively similar in that both the Sala-i-Martin *et al.* and Fernández *et al.* datasets yield similar amounts of parameter heterogeneity due to differences in initial income and human capital. The results from the Fernández *et al.* dataset are presented here since that is a well-known cross-country growth dataset which has not been tested for multiple regimes due to initial conditions using Bayesian methods.

⁸ IBMA is discussed in the next section.

The model presented below is based on Fernández *et al.* (2001), Begun and Eicher (2008), and Begun (2008). Consider a linear regression with a dependent variable y (the growth rate of GDP per capita), n independent observations, and a $n \times k$ design matrix X . The Bayesian problem is to search for the “best” models which take the form

$$y = \alpha \zeta_n + \beta_j X_j + \varepsilon, \quad (1)$$

where ζ_n is a n -dimensional vector of ones, X_j is an $n \times k_j$ submatrix of regressors in X , and β_j represents a vector of regression coefficients with $\beta_j \in \mathfrak{R}^{k_j}$ ($0 \leq k_j \leq k$). Given a set of possible models $M = \{M_1, M_2, \dots, M_v\}$ where $v = 2^k$, and data D , BMA ranks each model according to its posterior probability, $p(M_i | D)$, which equals the model’s share of the total posterior mass.

Therefore,

$$p(M_i | D) = \frac{p(D | M_i) p(M_i)}{\sum_{j=1}^v p(D | M_j) p(M_j)} \quad (2)$$

where

$$p(D | M_i) = \int p(D | \delta_i, M_i) p(\delta_i | M_i) d\delta_i \quad (3)$$

and $\delta_i, \delta_i \sim [\beta_i, \sigma^2]$, is a vector of parameters from model M_i . A weighted sum of the posterior probabilities of all models which contain a parameter of interest Θ can then be used to rank that parameter. Thus,

$$p(\Theta | D) = \sum_{i=1}^{2^k} p(\Theta | D, M_i) p(M_i | D) \quad (4)$$

yields the posterior inclusion probability for any Θ . Raftery (1993) gives the posterior means and variances

$$E[\beta | D] = \sum_{i=0}^v \hat{\beta}_i pr(M_i | D), \quad (5)$$

$$Var[\beta | D] = \sum_{i=0}^v (Var[\beta | D, M_i] + \hat{\beta}_i^2) pr(M_i | D) - E[\beta | D]^2. \quad (6)$$

One of the main difficulties in implementing BMA is choosing prior model probabilities and parameter distributions. Since there is no useful information on prior model probabilities in the present study, this paper assumes a uniform distribution over the model space (see, e.g., Raftery *et al.*, 1997, and Hoeting *et al.*, 1999). Thus, the prior probability of each model is 2^{-k} and each model has an equal chance of being included. Choosing prior parameter distributions, however, can sometimes be more difficult. The software utilized in this study uses the unit information prior (UIP), a diffuse and commonly used prior with ties to frequentist statistics. The UIP is a multivariate normal prior with the maximum likelihood estimate as its mean, and whose variance is given by the expected information matrix for one observation (Kass and Wasserman, 1995; Raftery, 1999).

V. Iterative Bayesian Model Averaging

After allowing for parameter heterogeneity in the 41 independent variables from Fernández *et al.* (2001), there are up to 80 candidate regressors in this study (which implies 2^{80} , or more than 1,200,000,000,000,000,000,000,000 possible models). Since the *bicreg* algorithm which is often used for running BMA can only analyze a maximum of 54 variables, this study uses Iterative Bayesian Model Averaging to test for model uncertainty.⁹ IBMA was initially developed for a genome application by Yeung, Baumgarner, and Raftery (2005) and later brought to economics in Eicher *et al.*'s 2007 study of cross-country growth heterogeneity.

⁹ Ley and Steel (2007) extended their MC³ algorithm to handle up to 104 regressors, and thus it is now possible to run traditional BMA on 80 variables. An advantage of IBMA is that it is extremely computationally efficient when faced with a large pool of potential regressors.

The main advantage of IBMA over traditional BMA is that it can process a greater number of candidate regressors by performing BMA repeatedly on a reduced set of variables. IBMA works by taking the first group of w variables which are most highly correlated with the dependent variable for the initial BMA run ($w = 41$ in this study). The subset of variables with very low posterior inclusion probabilities (usually less than one percent) are then substituted for with the subset of candidate regressors that 1) did not appear in the initial BMA run; and 2) are most highly correlated with the dependent variable, and BMA is run a second time. The process repeats in this fashion until all candidate variables have appeared in one or more BMA runs.

VI. Estimation

This study uses the well-known Fernández, Ley, and Steel (2001) dataset in its IBMA analysis. The dataset contains cross-sectional data for 41 variables which cover various macroeconomic, political, geographical, demographic, linguistic, historical, and religious characteristics of 72 countries in both developed and developing nations. The dependent variable is average per capita GDP growth from 1960-1992.

The methodology used in this study follows Eicher *et al.* (2007) and Begun (2008) that employ IBMA to test for heterogeneous growth in certain groups of countries (OECD countries in the former study and East Asian countries in the latter). The present study uses interaction terms for countries with initial GDP per capita above the median level, human capital¹⁰ above the median level, and both initial income and human capital above the median in order to allow for parameter heterogeneity.

¹⁰ Since a unique, widely-agreed-upon measure of human capital does not exist, both primary schooling and public education spending as a percent of GDP are used as human capital proxies.

As previously mentioned, the use of Iterative Bayesian Model Averaging enables all of the 41 candidate regressors from Fernández *et al.* (2007) to be tested for heterogeneity based on varying initial conditions. The main disadvantage to this approach is that this increases the number of potential regressors in the study to up to 80 once interaction terms are allowed for, and thus due to the relatively small number of countries in the dataset the sample cannot be split into endogenously determined subgroups using the regression tree procedure of Durlauf and Johnson (1995), or with the threshold-identifying procedure in Hansen (2000). Allowing for multiple subgroups would cause the number of candidate regressors to dwarf the number of countries, and testing for parameter heterogeneity in a small subset of countries (a possible outcome of Hansen’s procedure) would also not be feasible given the large number of candidate regressors since it would produce matrices of less than full column rank.

In order to test for heterogeneous growth in the sample, the model employed in this study starts with the standard cross-country growth model from equation (1) and then allows for parameter heterogeneity in a subset of countries. Incorporating parameter heterogeneity into the base model yields the standard interaction model:

$$y = \alpha\zeta_n + \beta_{1,j}X_j + \lambda\beta_{2,j}Z_j + \varepsilon, \quad (7)$$

where λ is an indicator function that takes the value of 1 if a country has high initial income, high initial human capital, or both high initial income and a high level of human capital depending on the specification, and Z is a submatrix of X that excludes regressors from the high income/human capital subsample that are either perfectly (or nearly perfectly) collinear, or that contain less than two non-zero observations.¹¹ The parameter estimates given in Tables 1-5 for the control group (e.g., countries with low initial income) are thus β_{1i} , whereas the parameter estimates for the

¹¹ See the notes to Tables 1-5 for a list of which variables were not tested for in the high income/high human capital subsamples.

treatment group (e.g., countries with high initial income) are given by the composite means

$$\tilde{\beta}_i = \beta_{1i} + \beta_{2i} \text{ [with } \text{var}(\tilde{\beta}_i) = \text{var}(\beta_{1i}) + \text{var}(\beta_{2i}) + 2\text{cov}(\beta_{1i}, \beta_{2i}) \text{]}.$$

VII. Results

The results are presented in Tables 1-5. Each table displays posterior means and standard deviations for the global sample, the low income and/or human capital subsample, and the high income and/or human capital subsample. Due to the large number of candidate regressors, and in order to focus on the important differences between samples, only variables which are deemed to be “effective” are presented in these tables. Raftery (1995) argues that a variable must have a posterior inclusion probability of greater than 50 percent in order to be considered effective. This roughly translates to a posterior mean to standard deviation ratio greater than one in absolute value, the threshold used for presentation in the tables.

Table 1 shows the results when using initial GDP per capita as a threshold. Perhaps the most striking thing is the difference in the set of effective regressors in the global sample versus each of the two subsamples. Of the 20 variables which are considered effective in the global sample, only 10 are found to be effective for countries with initially low incomes and only 9 are found to be effective for countries with high initial incomes. In addition, 9 of the effective variables from the low-income subsample and 11 from the high-income subsample are not effective in the main sample.

There is also a good deal of heterogeneity between the low-initial-income and high-initial-income groups. There are 9 regressors that are found to be effective for growth in one group of countries but not the other group. Interestingly, the negative coefficient on initial GDP per capita is nearly 40 percent larger in absolute value for countries with initially low incomes,

which suggests that there's been a greater amount of conditional convergence over the sample period for countries that were initially relatively poor. Another potentially important finding is that equipment investment has a much larger posterior mean for countries with low initial incomes (it is more than 55% larger than for countries with high initial incomes), which suggests that increasing equipment investment is especially important for stimulating growth in low-income countries which generally have a dearth of capital. Also, life expectancy appears to be much more important for growth in countries which were initially relatively poor since the posterior mean for life expectancy is 50% higher for countries with low initial income. It is important to note that of the 32 growth determinants found to be effective in the global sample or one of the two subsamples, only 7 are found to be effective in all three samples. Thus, policy prescriptions based on cross-country growth studies which assume homogeneity may be stressing the importance of factors which are not robustly related to long-run growth in certain types of countries, and ignoring other factors which may be crucial.

The evidence for heterogeneous growth due to differences in initial human capital, seen in Tables 2 and 3, is somewhat more mixed. There does seem to be support for heterogeneous growth due to differences in initial public education spending as nearly half of the effective variables in the subsamples are different from the global sample, and 7 of the effective variables in the subsamples are different. In particular, the religion variables show a great deal of heterogeneity as four of them (Fraction Buddhist, Fraction Muslim, Fraction Protestant, and Fraction Catholic) are found to be effective for growth in countries with low initial public education or high initial public education, but not both. The percentage of non-Orthodox Christians (Catholics or Protestants) and Muslims has a more positive effect on economic growth in countries with an initially high public education share (compared to the effect in countries

with low initial public education share), while the percentage of Buddhists has a more positive effect in countries with initially low public education. Fernandez *et al.* (2001) and Barro and McCleary (2003) also find evidence that religiosity impacts economic growth, though neither study addresses heterogeneous effects of religion on growth due to varying initial conditions.

The posterior means of several regressors also change sign between the low- and high-public-education regimes; interestingly, the Latin America and Sub-Saharan Africa dummy variables switch sign from negative to positive for the low- and high-public-education subsamples, respectively. This suggests that once other explanatory variables are controlled for, there is actually a positive effect on growth of being a Sub-Saharan or Latin American country when a relatively large percentage of initial GDP is spent on public education. This striking finding appears to be a novel one in the published literature. Many prior cross-country studies such as Barro (1991), Sala-i-Martin (1997), and Sala-i-Martin *et al.* (2004) find a negative coefficient for Sub-Saharan Africa and Latin America, but they assume a homogeneous growth process. Durlauf and Johnson (1995) run a robustness test where they allow for heterogeneous effects of the Latin American and Sub-Saharan African dummies due to differences in initial human capital, but 1) they group their countries by both high initial human capital/high initial output and low initial human capital/low initial output in their test, and 2) they do not report the results for these dummies.

For the primary schooling threshold the results are much more mundane. The global sample and the two subsamples all yield between 20 and 22 effective regressors, with only slight differences among which variables are effective in each sample. The posterior means are also very similar for the global sample, countries with a relatively low level of initial primary schooling, and countries with a relatively high level of initial primary schooling.

The results for using initial income and human capital as a combined threshold are reported in Tables 4 and 5. These results are similar to the earlier findings when using initial income or public education share of GDP as an individual threshold in that there tends to be a good deal of disparity between the effective variables in the global sample and subsamples. The number of effective variables also varies widely, as countries with low initial income/human capital have 17 effective regressors in each IBMA analysis, and countries with high initial income/human capital have 12 to 13 effective variables (compared to 20 in the global sample).¹² One result that stands out is that non-equipment investment appears to be an especially important driver of growth in countries with low initial incomes and/or low initial levels of human capital since the posterior means for non-equipment investment for countries with low initial GDP and/or human capital are two to three times larger than the posterior mean for non-equipment investment in the global sample.

As a whole, these results lend support to studies such as Durlauf and Johnson (1995), Hansen (2000), and Crespo Cuaresma (2002) that find that initial income and human capital are important threshold variables for heterogeneous growth. The findings, however, tend to run counter to Crespo Cuaresma and Doppelhofer's 2007 Bayesian Averaging of Thresholds (BAT) study which uses a subset of the Sala-i-Martin *et al.* (2004) dataset to conclude that initial income is not an important threshold variable.

VIII. Conclusion

This study allows for parameter heterogeneity and model uncertainty through use of Iterative Bayesian Model Averaging (IBMA) and supports a steady stream of empirical studies that find

¹² Note that dummy variables such as Sub-Saharan Africa cannot be considered effective for the high income/human capital subsamples since there are no countries in Sub-Saharan Africa that meet the criteria. See the notes to the tables for a list of regressors which were not tested for parameter heterogeneity.

evidence for heterogeneous economic growth across countries. This paper adds to the literature on country heterogeneity by using IBMA on a full suite of candidate regressors from a well-known cross-country dataset (Fernández, Ley, and Steel, 2001) to test for threshold effects in initial conditions. Specifically, it finds that initial income and public education share are important for determining a country's subsequent growth process.

A growing number of empirical studies over the past decade and a half (Durlauf and Johnson, 1995; Hansen, 2000; Masanjala and Papageorgiou, 2004; etc.) have found evidence to support parameter heterogeneity due to initial conditions. This has important implications for policy makers, as the “one-size-fits-all” approach for encouraging growth so often used in the past may be dubious. The main question now appears not to be whether differences in initial conditions lead to multiple regimes, but rather which initial conditions are robustly related to certain forms of heterogeneity in the growth process.

References

- Ardıç, O.P., 2006, “The Gap between the Rich and the Poor: Patters of Heterogeneity in the Cross-Country Data,” *Economic Modelling* 23: 538-555.
- Azariadis, C. and A. Drazen, 1990, “Threshold Externalities in Economic Development,” *Quarterly Journal of Economics* 105(2): 501-526.
- Barro, R. J., 1991, “Economic Growth in a Cross Section of Countries,” *Quarterly Journal of Economics* 106: 407-443.
- Barro, R. J., and R. M. McCleary, 2003, “Religion and Economic Growth Across Countries.” *American Sociological Review* 68: 760-781.
- Begun, J., 2009, “Different Growth for Different Folks? Reexamining East Asian Growth Determinants through Iterative Bayesian Model Averaging (IBMA),” Working Paper, Department of Economics, University of Washington.

- Begun, J. and T. S. Eicher, 2008, "In Search of an Environmental Kuznets Curve in Sulphur Dioxide Concentrations: A Bayesian Model Averaging Approach," *Environment and Development Economics* 13(6): 795-822.
- Brock, W. A. and S.N. Durlauf, 2001, "Growth Empirics and Reality," *World Bank Economic Review* 15(2): 229-272.
- Crespo Cuaresma, J., 2002, "Some Million Thresholds: Nonlinearity and Cross-Country Growth Regressions," Working Paper 0210, Department of Economics, University of Vienna.
- Crespo Cuaresma, J. and G. Doppelhofer, 2007, "Nonlinearities in Cross-Country Regressions: A Bayesian Averaging of Thresholds (BAT) Approach," *Journal of Macroeconomics* 29(3): 541-554.
- Draper, P., 1995, "Assessment and Propagation of Model Uncertainty" (with discussion), *Journal of the Royal Statistical Society (Series B)* 57: 45-97.
- Durlauf, S.N., 1993, "Nonergodic Economic Growth," *Review of Economic Studies* 60: 349-366.
- Durlauf, S.N. and P.A. Johnson, 1995, "Multiple Regimes and Cross-Country Growth Behavior," *Journal of Applied Econometrics* 10(4): 365-384.
- Durlauf, S.N., P. Johnson, and J. Temple, 2005, "Growth Econometrics," in *Handbook of Economic Growth*, P. Aghian and S.N. Durlauf (eds.), Amsterdam: North-Holland.
- Durlauf, S.N., A. Kourtellos, and A. Minkin, 2001, "The Local Solow Growth Model," *European Economic Review* 45: 928-940.
- Eicher, T. S., C. Papageorgiou, and O. Roehn, 2007, "Unraveling the Fortunes of the Fortunate: An Iterative Bayesian Model Averaging (IBMA) Approach," *Journal of Macroeconomics* 29: 494-514.
- Fernández, C., E. Ley, and M.F.J. Steel, 2001, "Model Uncertainty in Cross-Country Growth Regressions," *Journal of Applied Econometrics* 16: 563-576.
- Galor, O. and J. Zeria, 1993, "Income Distribution and Macroeconomics," *Review of Economic Studies* 60: 35-52.
- Hansen, B.E., 2000, "Sample Splitting and Threshold Estimation," *Econometrica* 68(3): 575-603.
- Hoeting, J., D. Madigan, A.E. Raftery, and C. Volinsky, 1999, "Bayesian Model Averaging: A Tutorial" (with discussion), *Statistical Science* 14: 382-401.

- Kass, R.E. and L. Wasserman, 1995, "A Reference Bayesian Test for Nested Hypotheses and Its Relationship to the Schwartz Criterion," *Journal of the American Statistical Association* 90: 928-934.
- Kormendi, R. C. and P. Meguire, 1985, "Macroeconomic Determinants of Growth: Cross-Country Evidence," *Journal of Monetary Economics* 16: 141-163.
- Leamer, E.E., 1978, *Specification Searches*, New York: Wiley.
- Leamer, E.E., 1983, "Let's Take the Con Out of Econometrics," *American Economic Review* 73: 31-43.
- Ley, E., and M.F.J. Steel, 2007, "Jointness in Bayesian Variable Selection with Applications to Growth Regression," *Journal of Macroeconomics* 29(3): 436-493.
- Liu, Z. and T. Stengos, 1999, "Non-Linearities in Cross-Country Growth Regressions: A Semiparametric Approach," *Journal of Applied Econometrics* 14: 527-538.
- Mankiw, N.G., D. Romer, and D.N. Weil, 1992, "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics* 107(2): 407-437.
- Masanjala, W.H. and C. Papageorgiou, 2004, "The Solow Model with CES Technology: Nonlinearities and Parameter Heterogeneity," *Journal of Applied Econometrics* 19: 171-201.
- Multon, B.R., 1991, "A Bayesian Approach to Regression Selection and Estimation with Application to a Price Index for Radio Services," *Journal of Econometrics* 49: 169-193.
- Palm, F.C. and A. Zellner, 1992, "To Combine or not to Combine? Issues of Combining Forecasts," *Journal of Forecasting* 11: 687-701.
- Papageorgiou, C., 2002, "Trade as a Threshold Variable for Multiple Regimes," *Economic Letters* 77: 85-91.
- Raftery, A.E., 1993, "Bayesian Model Selection in Structural Equation Models," in K.A. Bollen and J.S. Long (eds.), *Testing Structural Equation Models*, Beverly Hills: Sage, pp. 163-180.
- Raftery, A.E., 1995, "Bayesian Model Selection for Social Research," *Sociological Methodology* 25: 111-163.
- Raftery, A.E., D. Madigan, and J.A. Hoeting, 1997, "Bayesian Model Averaging for Linear Regression Models," *Journal of the American Statistical Association* 92: 179-191.
- Raftery, A.E., 1999, "Bayes Factors and BIC: Comment on Weakliem," *Sociological Methods and Research* 27: 411-427.

Sala-i-Martin, X.,G., 1997, "I Just Ran Two Million Regressions," *American Economic Review, Papers and Proceedings* 87: 178-183.

Sala-i-Martin, X., G. Doppelhofer, and R. I. Miller, 2004, "Determinants of Long-Term Growth: A Bayesian Model Averaging of Classical Estimates (BACE) Approach," *American Economic Review* 94: 813-835.

Yeung, K.Y., R.E. Baumgarner, and A.E. Raftery, 2005, "Bayesian Model Averaging: Development of an Improved Multi-Class, Gene Selection and Classification Tool for Microarray Data," *Bioinformatics* 21: 2394-2402.

Table 1
Effective Growth Determinants – Initial GDP per Capita Threshold

	Global Sample		Low Initial GDP per Capita		High Initial GDP per Capita	
	Posterior Mean	Posterior S.D.	Posterior Mean	Posterior S.D.	Posterior Mean	Posterior S.D.
Life expectancy in 1960	0.0009***	0.0002	0.0012***	0.0002	0.0008***	0.0003
GDP per capita in 1960	-0.0180***	0.0022	-0.0222***	0.0021	-0.0160***	0.0020
Sub-Saharan Africa	-0.0221***	0.0043	-0.0249***	0.0029		
Fraction Confucius	0.0737***	0.0102			0.1105***	0.0150
Rule of law	0.0125***	0.0037				
Equipment investment	0.1479***	0.0357	0.2650***	0.0362	0.1464***	0.0320
Fraction Hindu	-0.1080***	0.0202	-0.0579***	0.0113		
Size of labor force	0.0037***	0.0007	0.0014***	0.0004	0.0014***	0.0004
Higher education enrollment	-0.1205***	0.0294				
Ethnolinguistic fragmentation	0.0155***	0.0041				
Mining (share of GDP)	0.0330***	0.0117				
Latin America	-0.0130***	0.0047	-0.0111***	0.0027	-0.0111***	0.0027
Spanish colony	0.0137***	0.0049	0.0100***	0.0027	0.0100***	0.0027
British colony	0.0074***	0.0031				
French colony	0.0108***	0.0040	0.0095***	0.0022		
Primary schooling in 1960	0.0203**	0.0089			-0.0313***	0.0125
Civil liberties	-0.0023**	0.0011				
Outward orientation	-0.0033	0.0020				
Fraction speaking English	-0.0067	0.0044				
Non-equipment investment	0.0309	0.0210	0.0764***	0.0111	0.0764***	0.0111
Fraction Catholic			-0.0167***	0.0033	-0.0167***	0.0033
Degree of capitalism			0.0026***	0.0005	0.0026***	0.0005
Political rights			-0.0028***	0.0005	-0.0028***	0.0005
Primary exports, 1970			-0.0185***	0.0037	-0.0185***	0.0037
Exchange rate distortions			0.0001***	0.00002	-0.0002***	0.00004
War dummy			-0.0045***	0.0012	-0.0045***	0.0012
Fraction Buddhist			0.0089***	0.0032	0.0090***	0.0032
Revolutions and coups			0.0095***	0.0035		
Black market premium			-0.0036**	0.0018	0.0311***	0.0081
Population growth					-0.6052***	0.1636
Public education share					-0.2172**	0.0901
Workers/population (log)					-0.0274**	0.0116

Note: A variable is considered effective in this table if the ratio of its posterior mean to its posterior standard deviation is greater than one in absolute value. “Size of labor force” was multiplied by 10,000. Sub-Saharan Africa, Fraction Hindu, and French colony were not tested for heterogeneity since their interaction terms were either perfectly (or extremely highly) collinear or contained less than two non-zero observations. *, **, *** indicate posterior mean to posterior standard deviation ratios of greater than 1.70, 2.04, and 2.45 in absolute value, respectively (which roughly corresponds to 90, 95, and 99 percent confidence levels in frequentist hypothesis testing).

Table 2
Effective Growth Determinants – Public Education Share Threshold

	Global Sample		Low Public Education Share		High Public Education Share	
	Posterior Mean	Posterior S.D.	Posterior Mean	Posterior S.D.	Posterior Mean	Posterior S.D.
Life expectancy in 1960	0.0009***	0.0002			0.0017***	0.0002
GDP per capita in 1960	-0.0180***	0.0022	-0.0127***	0.0015	-0.0132***	0.0019
Sub-Saharan Africa	-0.0221***	0.0043	-0.0197***	0.0028	0.0154***	0.0055
Fraction Confucius	0.0737***	0.0102	0.0961***	0.0084	0.0961***	0.0084
Rule of law	0.0125***	0.0037	0.0184***	0.0031	0.0184***	0.0031
Equipment investment	0.1479***	0.0357			0.1788***	0.0298
Fraction Hindu	-0.1080***	0.0202	-0.0699***	0.0117	0.1140***	0.0397
Size of labor force	0.0037***	0.0007	0.0016***	0.0005	0.0016***	0.0005
Higher education enrollment	-0.1205***	0.0294				
Ethnolinguistic fragmentation	0.0155***	0.0041				
Mining (share of GDP)	0.0330***	0.0117				
Latin America	-0.0130***	0.0047	-0.0070***	0.0024	0.0093***	0.0033
Spanish colony	0.0137***	0.0049				
British colony	0.0074***	0.0031				
French colony	0.0108***	0.0040	-0.0060***	0.0021	0.0121***	0.0029
Primary schooling in 1960	0.0203**	0.0089				
Civil liberties	-0.0023**	0.0011	0.0049***	0.0015	0.0049***	0.0015
Outward orientation	-0.0033	0.0020				
Fraction speaking English	-0.0067	0.0044	-0.0118***	0.0029	-0.0118***	0.0029
Non-equipment investment	0.0309	0.0210	0.0502***	0.0113	0.0502***	0.0113
Age			-0.0001***	0.00002	-0.0001***	0.00002
Fraction Buddhist			0.0120***	0.0042		
Workers/population (log)			-0.0167***	0.0047	-0.0167***	0.0047
Black market premium			-0.0138***	0.0021	-0.0138***	0.0021
Degree of capitalism			0.0051***	0.0006	0.0048***	0.0006
Years economy open					-0.0196***	0.0047
Fraction Muslim					0.0405***	0.0077
Fraction Protestant			-0.0221***	0.0053		
Political rights			-0.0077***	0.0012	-0.0057***	0.0014
Foreign language spoken (%)			-0.0092***	0.0025	-0.0099***	0.0025
Fraction Catholic					0.0204***	0.0057

Note: A variable is considered effective in this table if the ratio of its posterior mean to its posterior standard deviation is greater than one in absolute value. “Size of labor force” was multiplied by 10,000. The posterior means for “Size of labor force” were multiplied by 10,000. Fraction Jewish, Fraction Confucius, and Spanish colony were not tested for heterogeneity since their interaction terms were either perfectly (or extremely highly) collinear or contained less than two non-zero observations. *, **, *** indicate posterior mean to posterior standard deviation ratios of greater than 1.70, 2.04, and 2.45 in absolute value, respectively (which roughly corresponds to 90, 95, and 99 percent confidence levels in frequentist hypothesis testing).

Table 3
Effective Growth Determinants – Initial Primary Schooling Threshold

	Global Sample		Low Initial Primary Schooling		High Initial Primary Schooling	
	Posterior Mean	Posterior S.D.	Posterior Mean	Posterior S.D.	Posterior Mean	Posterior S.D.
Life expectancy in 1960	0.0009***	0.0002	0.0009***	0.0002	0.0003	0.0002
GDP per capita in 1960	-0.0180***	0.0022	-0.0140***	0.0021	-0.0140***	0.0021
Sub-Saharan Africa	-0.0221***	0.0043	-0.0167***	0.0028	-0.0167***	0.0028
Fraction Confucius	0.0737***	0.0102	0.0601***	0.0090	0.0601***	0.0090
Rule of law	0.0125***	0.0037				
Equipment investment	0.1479***	0.0357	0.3185***	0.0482	0.1140***	0.0343
Fraction Hindu	-0.1080***	0.0202	-0.1160***	0.0144	-0.1160***	0.0144
Size of labor force	0.0037***	0.0007	0.0040***	0.0006	0.0040***	0.0006
Higher education enrollment	-0.1205***	0.0294	-0.0952	0.0691	-0.0797***	0.0289
Ethnolinguistic fragmentation	0.0155***	0.0041	0.0121***	0.0032	0.0121***	0.0032
Mining (share of GDP)	0.0330***	0.0117				
Latin America	-0.0130***	0.0047	-0.0200***	0.0041	-0.0227***	0.0041
Spanish colony	0.0137***	0.0049	0.0239***	0.0053	0.0187***	0.0044
British colony	0.0074***	0.0031	0.0098***	0.0023	0.0098***	0.0023
French colony	0.0108***	0.0040	0.0137***	0.0027	0.0137***	0.0027
Primary schooling in 1960	0.0203**	0.0089				
Civil liberties	-0.0023**	0.0011	-0.0021***	0.0007	-0.0021***	0.0007
Outward orientation	-0.0033	0.0020	-0.0044***	0.0015	-0.0044***	0.0015
Fraction speaking English	-0.0067	0.0044	-0.0188***	0.0065	-0.0127***	0.0048
Non-equipment investment	0.0309	0.0210	0.0778***	0.0138	0.0778***	0.0138
Degree of capitalism			0.0017***	0.0006	0.0017***	0.0006
War dummy					-0.0087***	0.0025
Fraction Buddhist			0.0124***	0.0049	0.0086	0.0051
Population growth			-0.5415***	0.1486	-0.5415***	0.1486
Workers/population (log)					-0.0463***	0.0092

Note: A variable is considered effective in this table if the ratio of its posterior mean to its posterior standard deviation is greater than one in absolute value. “Size of labor force” was multiplied by 10,000. Sub-Saharan Africa, French colony, and Primary schooling were not tested for heterogeneity since their interaction terms were either perfectly (or extremely highly) collinear or contained less than two non-zero observations. *, **, *** indicate posterior mean to posterior standard deviation ratios of greater than 1.70, 2.04, and 2.45 in absolute value, respectively (which roughly corresponds to 90, 95, and 99 percent confidence levels in frequentist hypothesis testing).

Table 4
Effective Growth Determinants – High Initial GDP
and Public Education Share Threshold

	Global Sample		Low Initial GDP and/or Public Education Share		High Initial GDP and Public Education Share	
	Posterior Mean	Posterior S.D.	Posterior Mean	Posterior S.D.	Posterior Mean	Posterior S.D.
Life expectancy in 1960	0.0009***	0.0002	0.0008***	0.0002	0.0008***	0.0002
GDP per capita in 1960	-0.0180***	0.0022	-0.0130***	0.0025	-0.0129***	0.0025
Sub-Saharan Africa	-0.0221***	0.0043	-0.0189***	0.0035		
Fraction Confucius	0.0737***	0.0102	0.0583***	0.0089		
Rule of law	0.0125***	0.0037	0.0076	0.0049	0.0076	0.0050
Equipment investment	0.1479***	0.0357	0.1031	0.0646	0.1331***	0.0427
Fraction Hindu	-0.1080***	0.0202	-0.0754***	0.0193		
Size of labor force	0.0037***	0.0007	0.0280***	0.0082	0.0278***	0.0083
Higher education enrollment	-0.1205***	0.0294	-0.0998***	0.0350	-0.0998***	0.0350
Ethnolinguistic fragmentation	0.0155***	0.0041				
Mining (share of GDP)	0.0330***	0.0117	0.0449***	0.0135	0.0449***	0.0135
Latin America	-0.0130***	0.0047	-0.0114***	0.0042	-0.0114***	0.0044
Spanish colony	0.0137***	0.0049				
British colony	0.0074***	0.0031				
French colony	0.0108***	0.0040				
Primary schooling in 1960	0.0203**	0.0089				
Civil liberties	-0.0023**	0.0011				
Outward orientation	-0.0033	0.0020				
Fraction speaking English	-0.0067	0.0044				
Non-equipment investment	0.0309	0.0210	0.0619***	0.0173		
Age			-0.00004	0.00003	-0.00004	0.00003
Degree of capitalism			0.0026***	0.0007	0.0026***	0.0007
Black market premium			-0.0089***	0.0027	-0.0089***	0.0027
Fraction Protestant			-0.0268***	0.0069		
Public education share			0.2387*	0.1239	0.2387*	0.1239

Note: A variable is considered effective in this table if the ratio of its posterior mean to its posterior standard deviation is greater than one in absolute value. “Size of labor force” was multiplied by 10,000. Fraction Confucius, Sub-Saharan Africa, Fraction Hindu, French colony, and Fraction Buddhist were not tested for heterogeneity since their interaction terms were either perfectly (or extremely highly) collinear or contained less than two non-zero observations. *, **, *** indicate posterior mean to posterior standard deviation ratios of greater than 1.70, 2.04, and 2.45 in absolute value, respectively (which roughly corresponds to 90, 95, and 99 percent confidence levels in frequentist hypothesis testing).

Table 5
Effective Growth Determinants – High Initial GDP
and Primary Schooling Threshold

	Global Sample		Low Initial GDP and/or Primary Schooling		High Initial GDP and Primary Schooling	
	Posterior Mean	Posterior S.D.	Posterior Mean	Posterior S.D.	Posterior Mean	Posterior S.D.
Life expectancy in 1960	0.0009***	0.0002	0.0010***	0.0002	0.0010***	0.0002
GDP per capita in 1960	-0.0180***	0.0022	-0.0160***	0.0020	-0.0160***	0.0020
Sub-Saharan Africa	-0.0221***	0.0043	-0.0187***	0.0030		
Fraction Confucius	0.0737***	0.0102	0.0459***	0.0088		
Rule of law	0.0125***	0.0037				
Equipment investment	0.1479***	0.0357	0.2405***	0.0493	0.1497***	0.0504
Fraction Hindu	-0.1080***	0.0202	-0.0883***	0.0151		
Size of labor force	0.0037***	0.0007	0.0032***	0.0007	0.0032***	0.0007
Higher education enrollment	-0.1205***	0.0294	-0.1175***	0.0267	-0.1175***	0.0267
Ethnolinguistic fragmentation	0.0155***	0.0041				
Mining (share of GDP)	0.0330***	0.0117				
Latin America	-0.0130***	0.0047	-0.0201***	0.0034	-0.0201***	0.0034
Spanish colony	0.0137***	0.0049	0.0133***	0.0039	0.0118***	0.0042
British colony	0.0074***	0.0031				
French colony	0.0108***	0.0040	0.0069***	0.0027		
Primary schooling in 1960	0.0203**	0.0089				
Civil liberties	-0.0023**	0.0011				
Outward orientation	-0.0033	0.0020				
Fraction speaking English	-0.0067	0.0044	-0.0141	0.0090		
Non-equipment investment	0.0309	0.0210	0.0966***	0.0170	-0.0397	0.0261
Age			-0.0001***	0.00002	-0.0001***	0.00002
Degree of capitalism			0.0026***	0.0006	0.0026***	0.0007
Black market premium			-0.0062**	0.0026	-0.0062**	0.0026
Political rights			-0.0019***	0.0008	-0.0019***	0.0008
Workers/population (log)					-0.0239***	0.0056

Note: A variable is considered effective in this table if the ratio of its posterior mean to its posterior standard deviation is greater than one in absolute value. “Size of labor force” was multiplied by 10,000. Fraction Confucius, Sub-Saharan Africa, Fraction Hindu, French colony, and Fraction Buddhist were not tested for heterogeneity since their interaction terms were either perfectly (or extremely highly) collinear or contained less than two non-zero observations. *, **, *** indicate posterior mean to posterior standard deviation ratios of greater than 1.70, 2.04, and 2.45 in absolute value, respectively (which roughly corresponds to 90, 95, and 99 percent confidence levels in frequentist hypothesis testing).

APPENDIX

Table A-1: Descriptive Statistics from the Global Sample

Variable	Mean	Standard Deviation
Average GDP per capita growth (1960-1992)	0.021	0.018
Primary schooling in 1960	0.795	0.246
Life expectancy	56.581	11.448
GDP per capita in 1960	7.492	0.885
Mining (share of GDP)	0.045	0.077
Degree of capitalism	3.542	1.266
Years open	0.439	0.355
Fraction speaking English	0.076	0.239
Fraction speaking a foreign language	0.374	0.422
Real exchange rate distortions	121.708	41.001
Equipment investment	0.044	0.035
Non-equipment investment	0.149	0.055
Standard deviation of black market premium	45.596	95.802
Outward orientation	0.389	0.491
Black market premium	0.157	0.291
Area	972.917	2051.976
Latin America dummy	0.278	0.451
Sub-Saharan Africa dummy	0.208	0.409
Higher education enrollment in 1960	0.043	0.052
Public education share	0.025	0.009
Revolutions and coups	0.182	0.238
War dummy	0.403	0.494
Political rights	3.451	1.896
Civil liberties	3.466	1.712
Absolute latitude	25.733	17.250
Age	23.708	37.307
British colony	0.319	0.470
Fraction Buddhist	0.056	0.184
Fraction Catholic	0.422	0.397
Fraction Confucius	0.019	0.087
Ethnolinguistic fragmentation	0.371	0.296
French colony	0.125	0.333
Fraction Hindu	0.018	0.101
Fraction Jewish	0.013	0.097
Fraction Muslim	0.148	0.295
Primary Exports	0.673	0.299
Fraction Protestant	0.173	0.252
Rule of law	0.551	0.335
Spanish colony	0.222	0.419
Population growth	0.020	0.010
Workers to population ratio (log)	-0.954	0.189
Size of the labor force	9305.376	24906.056

Note: Lower values imply greater political rights and civil liberties.